

# METHOD OF FABRICATING MOS TRANSISTOR USING TOTAL GATE SILICIDATION PROCESS

## CROSS-REFERENCE TO RELATED APPLICATION

5           This application claims the benefit of Korean Patent Application No. 2003-46983 filed on July 10, 2003, the contents of which are hereby incorporated herein by reference in their entirety.

## BACKGROUND OF THE INVENTION

### 10           1. Field of the Invention

          The present invention relates to a method of fabricating semiconductor devices, and more particularly to a method of fabricating a MOS transistor using a total gate silicidation process.

### 15           2. Description of the Related Art

          MOS transistors are widely in use as discrete switching devices in semiconductor devices. With the increase in integration of semiconductor devices, the size of the MOS transistor has been reduced. As a result, the reduction in the channel length of the MOS-transistor causes a short channel effect.

20           The junction depth of the source/drain region of the MOS transistor is required to be reduced to improve the short channel effect. The reduction of the junction depth of the source/drain region increases the sheet resistance of the

source/drain region, which results in a decreased ability to provide an appropriate driving current required in a high performance transistor. Therefore, a self-aligned silicide process is widely used in order to decrease the sheet resistance of the source/drain region, and to reduce the sheet resistance of the gate electrode.

5         The self-aligned silicide process has been developed and widely used to form a silicide layer at the source/drain region and on top of the polysilicon of the gate region at the same time. In accordance with this process, the thickness of the silicide layer formed at the source/drain region should be smaller than the junction depth.

10         Also, with the MOS transistor scaled down, the thickness ( $T_{ox}$ ) of the gate insulating layer has been reduced to avoid the short channel effect and improve the current drivability. The reduction of the thickness of the gate insulating layer increases the capacitance of the gate insulating layer so as to improve the current drivability.

15         However, as a result of the reduction of the thickness of the gate insulating layer, poly depletion occurs in the MOS transistor, in which a polysilicon layer is used as a gate electrode. The poly depletion layer increases the electrical equivalent thickness of the gate insulating layer, and decreases the drive current.

20         As a method of solving the poly depletion problem, use of a metal gate has been widely studied. However, the metal gate has a drawback in the form of a difficulty in controlling threshold voltage of a transistor. Specifically, the threshold voltages of the NMOS transistor region and the PMOS transistor region should be controlled in order to employ the metal gate in a CMOS transistor. However, the

use of a different metal gate in each MOS transistor makes the fabrication processes too complicated to be practicably employed.

One method for solving the problems of poly depletion and the metal gate is taught by Tavel et. al. in a research paper entitled, "Totally Silicided (CoSi<sub>2</sub>) Polysilicon: a novel approach to very low-resistive gate ( $\sim 2\Omega/\square$ ) without metal CMP nor etching," (IEDM 2001, p.825-828, December of 2001).

However, the method disclosed in the above paper also has a drawback in that a threshold voltage ( $V_{th}$ ) is increased due to a decrease of the absolute value of a flat band voltage ( $V_{fb}$ ) since a midgap material is used as a metal gate, and that the fabrication processes are complicated since the silicide layer of the source/drain region and the gate silicide layer are formed respectively using separate processes. Specifically, in order to form the silicide layer of the source/drain region and the gate silicide layer using separate processes, a chemical mechanical polishing process (CMP) is required.

In order to concurrently form the silicide layer of the source/drain region and the gate silicide layer using the total gate silicide process, it is necessary to reduce the gate height. However, when the gate height is even slightly reduced, ions are injected into a channel region during the ion implantation process into the source/drain regions, which causes a problem that leakage current of the MOS transistor is increased.

As a result, it is necessary to optimize the fabrication method of the MOS transistor being capable of solving the poly depletion problem.

## **SUMMARY OF THE INVENTION**

The present invention provides a method of fabricating a MOS transistor and a CMOS transistor in which a silicide layer of source/drain regions and a gate silicide layer are formed concurrently. In accordance with the invention, ions are prevented from being injected into a channel region during an implantation of source/drain ions.

The present invention also provides a method of fabricating a MOS transistor and a CMOS transistor capable of maintaining a low threshold voltage with use of a metal gate.

The present invention is directed to a method of fabricating a MOS transistor using a total gate silicidation process. The method includes forming an insulated gate pattern on a semiconductor substrate. The insulated gate pattern may include a silicon pattern and a sacrificial layer pattern, which are sequentially stacked. Spacers covering sidewalls of the gate pattern are formed, and impurity ions are injected into the semiconductor substrate using the spacers and the gate pattern as ion injection masks to form source/drain regions. The sacrificial layer pattern on the semiconductor substrate having the source/drain regions is removed to expose the silicon pattern. The exposed silicon pattern is converted into a gate silicide layer, and source/drain silicide layers are concurrently selectively formed at surfaces of the source/drain regions.

The formation of the insulated gate pattern may include sequentially forming a gate insulating layer and a silicon layer on the semiconductor substrate. A sacrificial layer is formed on the semiconductor substrate having the silicon layer.

The sacrificial layer and the silicon layer are patterned sequentially to form the silicon pattern and the sacrificial layer pattern.

Before or after forming the sacrificial layer, the silicon layer may be doped with impurities to control a threshold voltage. The impurities are N-type or P-type according to types of the MOS transistor. That is, in the case of fabricating an NMOS transistor, the silicon layer is doped with N-type impurities, and in the case of fabricating a PMOS transistor, the silicon layer is doped with P-type impurities. Further, before doping the silicon layer with the impurities, a buffer layer may be formed on the silicon layer. The buffer layer may be used to prevent the impurities from penetrating into the semiconductor substrate under the gate insulating layer in a technology of implanting impurity ions using high energy. In this case, the buffer layer is etched to form a buffer layer pattern before patterning the silicon layer, and the buffer layer pattern is removed along with the sacrificial layer pattern to expose the silicon pattern.

Before forming the spacers, impurity ions may be injected into the semiconductor substrate using the gate pattern as an ion injection mask to form an LDD and a halo in order to prevent a short channel effect and a hot carrier effect. Further, in order to prevent a short channel effect, before forming the LDD and the halo, offset spacers may be formed. The offset spacers prevent the occurrence of the short channel effect by increasing the channel length.

The silicide layers may be formed by forming a metal layer on the semiconductor substrate having the exposed silicon pattern, and annealing the metal layer until the exposed silicon pattern is silicided. After the silicide layers

are formed, the unreacted portion of the metal layer remaining on the spacers is removed in order to electrically insulate the gate silicide layer and the source/drain silicide layer.

In one embodiment, the semiconductor substrate is a single crystal silicon substrate, a silicon-on-insulator (SOI) substrate, or a strained silicon substrate.

The spacers can be formed of an insulating layer having an etch selectivity with respect to the sacrificial layer pattern. The insulating layer having the etch selectivity is composed of multiple layers including a silicon oxide layer and a silicon nitride layer. The metal layer can include at least one metal layer or alloy layer including metals selected from the group consisting of Ni, Co, W, and Ti. The metal layer can be a nickel layer or a nickel alloy layer.

In one embodiment, a selective epitaxial growth layer is formed on the source/drain regions before removing the sacrificial layer pattern. In accordance with another aspect, the present invention provides a method of fabricating a CMOS transistor. The method includes defining an NMOS transistor region and a PMOS transistor region on a predetermined portion of a semiconductor substrate, and forming insulated gate patterns on the NMOS transistor region and the PMOS transistor region. Each of the insulated gate patterns may include a silicon pattern and a sacrificial layer pattern, which are sequentially stacked. Spacers covering the sidewalls of the gate patterns are formed, and impurity ions are injected into the NMOS transistor region and the PMOS transistor region using the gate patterns and the spacers as ion injection masks to form source/drain regions. N-type source/drain regions are formed in the NMOS transistor region, and P-type

source/drain regions are formed in the PMOS transistor region. By removing the sacrificial layer patterns on the semiconductor substrate having the source/drain regions, the silicon patterns are exposed, and the exposed silicon patterns are converted into gate silicide layers, and source/drain silicide layers are concurrently selectively formed at the surface of the source/drain regions.

The formation of the gate patterns can include sequentially forming a gate insulating layer and a silicon layer on the semiconductor substrate. A sacrificial layer is formed on the semiconductor substrate having the silicon layer, and the sacrificial layer and the silicon layer are patterned sequentially to form the silicon patterns and the sacrificial layer patterns.

Before or after forming the sacrificial layer, the silicon layer on the upper portion of the NMOS transistor region and the PMOS transistor region may be doped with impurities to control threshold voltages. In this case, the silicon layer on the upper portion of the NMOS transistor region may be doped with N-type impurities, and the silicon layer on the upper portion of the PMOS transistor region may be doped with P-type impurities. Further, before doping the silicon layer, a buffer layer may be formed on top of the silicon layer. The buffer layer can be etched to form buffer layer patterns before patterning the silicon layer, and the buffer layer patterns can be removed along with the sacrificial layer patterns in order to expose the silicon patterns.

Before forming the spacers, impurity ions may be injected into the NMOS transistor region and the PMOS transistor region of the semiconductor substrate

having the gate patterns to form LDDs and halos. Further, before forming the LDDs and the halos, there may be formed offset spacers.

In one embodiment, the semiconductor substrate is a single crystal silicon substrate, a silicon-on-insulator (SOI) substrate, or a strained silicon substrate.

5 The spacers can be formed of an insulating layer having an etch selectivity with respect to the sacrificial layer patterns. The insulating layer having the etch selectivity is composed of multiple layers including a silicon oxide layer and a silicon nitride layer.

In one embodiment, converting the exposed silicon patterns into gate silicide layers and concurrently forming source/drain silicide layers selectively, comprises: i) forming a metal layer on the semiconductor substrate having the exposed silicon patterns; ii) annealing the metal layer until the exposed silicon patterns are totally silicided; and iii) removing the unreacted metal layer remaining on the spacers. The metal layer can be at least one metal layer or an alloy layer of a metal selected from the group consisting of Ni, Co, W, and Ti.

15 The metal layer can be a nickel layer or a nickel alloy layer.

In one embodiment, a selective epitaxial growth layer is formed on the source/drain regions before removing the sacrificial layer pattern.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

20 The foregoing and other features and advantages of the invention will be apparent from the more particular description of an embodiment of the invention, as illustrated in the accompanying drawing. The drawing is not necessarily to scale, emphasis instead being placed upon illustrating the principles of the invention.



FIG. 1 is a process flow chart illustrating a method of fabricating a CMOS transistor according to one embodiment of the present invention.

FIGs. 2 to 7 are sectional views illustrating a method of fabricating a CMOS transistor according to one embodiment of the present invention.

5

### **DETAILED DESCRIPTION OF THE INVENTION**

The present invention will now be described more fully hereinafter with reference to the accompanying drawings, in which preferred embodiments of the invention are shown. In the drawings, the thicknesses of layers and regions are  
10 exaggerated for clarity.

FIG. 1 is a process flow chart illustrating a method of fabricating a CMOS transistor according to one embodiment of the present invention, and FIGs. 2 to 7 are sectional views illustrating a method of fabricating a CMOS transistor according to one embodiment of the present invention. In FIGs. 2 to 7, the reference  
15 numeral A represents an NMOS transistor region, and the reference numeral B represents a PMOS transistor region.

Referring to FIGs. 1 and 2, a device isolation layer 53 is formed in a predetermined portion of a semiconductor substrate 51 to define active regions. The semiconductor substrate 51 may be a single crystal silicon substrate, a silicon-on-insulator (SOI) substrate, or a strained silicon substrate. In the case of a  
20 CMOS transistor, an NMOS transistor region A and a PMOS transistor region B are defined (step 1 of FIG. 1). The NMOS transistor region A and the PMOS transistor region B may be defined by forming wells, but a fully depleted SOI (FDSOI)

substrate does not require the formation of wells. In the illustrated embodiment, P-type impurities are injected to form wells in the NMOS transistor region A, and N-type impurities are injected to form wells in the PMOS transistor region B.

5 A gate insulating layer 55 and a silicon layer 57 are sequentially formed on the semiconductor substrate having the active regions (step 3 of FIG. 1). The gate insulating layer 55 is preferably formed of a silicon oxide ( $\text{SiO}_2$ ) layer.

The silicon layer 57 can be formed of a polysilicon layer or an amorphous silicon layer, and in order to fabricate a CMOS transistor, it is preferable to form an undoped silicon layer. Further, in order to reduce the gate height, the silicon layer  
10 57 is preferably formed with a thickness of 100Å to 600Å.

Referring to FIGs. 1 and 3, the silicon layer 57 may be doped with impurities to control a threshold voltage ( $V_{th}$ ) (step 5 of FIG. 1). When the silicon layer 57 is fully converted into a metal silicide layer, the impurities function to control the surface workfunction of the metal silicide layer to control the threshold  
15 voltage.

The doping of the impurities can be performed using a plasma doping technology or a high energy ion implantation technology. That is, by using a mask, the silicon layer 57 on top of the NMOS transistor region A is doped with N-type impurities, and the silicon layer 57 on top of the PMOS transistor region B is doped  
20 with P-type impurities.

Herein, the doping impurities are opposite types to the impurities in the well formed on the semiconductor substrate under the gate insulating layer 55. Therefore, the impurities injected to dope the silicon layer 57 should be prevented

from penetrating through the gate insulating layer 55. It is preferable to form a buffer layer 59 on the silicon layer 57 before doping with the impurities using the high energy ion implantation technology. The buffer layer 59 can be formed of SiN or SiON, and preferably, it is formed of a silicon nitride (SiN) layer. The buffer layer 59 is formed with an appropriate thickness with consideration of a penetration depth of implanted ions.

Referring to FIGs. 1 and 4, after doping the silicon layer 57 with the impurities, a sacrificial layer is formed on the buffer layer 59. The sacrificial layer is formed of SiN or SiON, and preferably, it is formed of a silicon nitride (SiN) layer.

The buffer layer 59 and the sacrificial layer function as masks in a subsequent source/drain ion implantation process, and prevent impurities from penetrating into gate channel regions. Therefore, in the case that the buffer layer 59 is thick enough, it is not necessary to form the sacrificial layer. In the case that no sacrificial layer is formed, the buffer layer 59 will be referred to as a sacrificial layer.

The sacrificial layer, the buffer layer 59, and the silicon layer 57 in the NMOS transistor region A and the PMOS transistor region B are patterned sequentially to form gate patterns 61a including silicon patterns 57a, buffer layer patterns 59a, and the sacrificial layer patterns 61 (step 7 of FIG. 1). Herein, the gate insulating layer 55 can also be patterned so as to form gate insulating layer patterns 55a.

Referring to FIGs. 1 and 5, using the gate patterns 61a as ion injection masks, impurities ions are injected into the NMOS transistor region A and the

PMOS transistor region B on the semiconductor substrate 51, so as to form lightly-doped drains (LDDs) 67, and halos 69 (step 9 of FIG. 1).

The LDDs 67 are formed by injecting N-type impurities into the NMOS transistor region A and P-type impurities into the PMOS transistor region B. To do this, anion injection mask other than that for the gate patterns 61a is formed. That is, by using an additional mask during the formation of LDDs in the NMOS transistor region A, N-type impurities can be prevented from being injected into the PMOS transistor region B.

The formation technology of the LDD is widely used in order to prevent a hot carrier effect. The halos 69 are used in order to avoid a short channel effect. The halos are formed by injecting impurities having an opposite type of that of the impurities injected to form the LDD. The use of an additional mask is the same as the case of forming the LDD.

In order to avoid the short channel effect by increasing the channel length, there are formed offset spacers 63 covering the sidewalls of the gate patterns 61a before forming the LDDs 67 and the halos 69. The offset spacers 63 are formed of an insulating layer having an etch selectivity with respect to the buffer layer patterns 59a and the sacrificial layer patterns 61. In one embodiment, the offset spacers 63 are formed of a silicon oxide ( $\text{SiO}_2$ ) layer.

After the LDDs 67 and the halos 69 are formed, a conformal spacer insulating layer is formed on the semiconductor substrate having the gate patterns 61a. The spacer insulating layer is formed of an insulating layer having an etch selectivity with respect to the buffer layer patterns 59a and the sacrificial layer

patterns 61. It may be formed of a silicon oxide ( $\text{SiO}_2$ ) layer, and more preferably, it may be formed of multiple layers composed of at least two layers, which include a silicon oxide layer and a silicon nitride layer. The silicon nitride layer functions to prevent etch damage on the silicon oxide layer while subsequent processes such as a cleaning process of a native oxide layer are carried out.

The spacer insulating layer over the NMOS transistor region A and the PMOS transistor region B of the semiconductor substrate 51 is etched back to form spacers 65 for covering the sidewalls of the offset spacers 63.

High dose impurities are injected into the NMOS transistor region A and the PMOS transistor region B of the semiconductor substrate having the spacers 65 to form source/drain regions 71 (step 11 of FIG. 1). In the illustrated embodiment, N-type impurities are injected into the NMOS transistor region A and P-type impurities are injected into the PMOS transistor region B.

As shown in FIG. 3, in the case that the silicon layer 57 is not doped with impurities to control threshold voltages, the silicon patterns 57a are made to be doped with impurities during the process of injecting the high dose impurities, so as to control the threshold voltages.

The buffer layer patterns 59a and the sacrificial layer patterns 61 function as masks against injected ions, so that the injected ions are prevented from passing through the gate insulating layer patterns 55a and penetrating into the channel region.

Referring to FIGs. 1 and 6, the sacrificial layer patterns 61 and the buffer layer patterns 59a are removed in order to expose the silicon patterns 57a in the

NMOS transistor region A and the PMOS transistor region B in which the source/drain regions 71 are formed.

Selective epitaxial growth (SEG) layers 73 may be formed in the NMOS transistor region A and the PMOS transistor region B having the source/drain regions 71 before removing the sacrificial layer patterns 61 and the buffer layer patterns 59a. The SEG layers 73 function to alleviate silicon consumption in the source/drain regions 71 during a process of forming silicide layers.

Referring to FIGs. 1 and 7, a metal layer is formed on the semiconductor substrate having the exposed silicon patterns 57a. As a result, the metal layer is formed on the exposed silicon patterns 57a. Further, the metal layer is formed on the surface of the source/drain regions 71.

The metal layer may be formed of one metal layer or at least one alloy layers selected from the group consisting of nickel (Ni) layer, cobalt (Co) layer, tungsten (W) layer, and titanium (Ti) layer. It may be formed of a nickel layer. The nickel layer may be a pure nickel layer or a nickel alloy layer. Herein, the nickel alloy layer may include tantalum (Ta), zirconium (Zr), titanium (Ti), hafnium (Hf), tungsten (W), cobalt (Co), platinum (Pt), palladium (Pd), vanadium (V), niobium (Nb) or a combination thereof.

The semiconductor substrate having the metal layer is annealed to form gate silicide layers 57b into which the exposed silicon patterns 57a are fully converted (step 13 of FIG. 1). In the case that the metal layer is a nickel layer, there can be formed a nickel silicide layer by annealing at a temperature of 400°C to 530°C in the presence of nitrogen. Herein, along with that, source/drain silicide

layers 73a are formed at the surface of the source/drain regions.

Even though the gate silicide layers 57b and the source/drain silicide layers 73a are concurrently formed, junctions of the source/drain regions 71 are not broken down since the silicon patterns 57a are thin,

5 Furthermore, in the case that the SEG layers 73 are formed on the source/drain regions 71, the SEG layers 73 react with the metal layer to form the silicide layer. Therefore, junction breakdown of the source/drain regions 71 can be avoided.

10 After the silicide layers 57b, 73a are formed, the unreacted portion of the metal layer remaining on the spacers 65 is removed such that the gate silicide layers 57b and the source/drain silicide layers 73a are electrically insulated. The unreacted portion of the metal layer remaining on top of the silicide layers 57b, 73a together is removed.

15 An interlayer insulating layer 75 is formed on the semiconductor substrate having the silicide layers 57b, 73a (step 15 of FIG. 1). In the case that the silicide layers 57b, 73a are nickel silicide layers, the interlayer insulating layer 75 is formed at a low temperature less than 550°C.

20 According to the present invention, a MOS transistor and a CMOS transistor can be provided through fabrication processes including a total gate silicidation process, in which source/drain silicide layers and a gate silicide layer are concurrently formed, and ions can be prevented from being injected into a channel region during an implantation of source/drain ions. Further, according to the present invention, there are provided a MOS transistor and a CMOS transistor

being capable of maintaining a low threshold voltage even with a use of a metal gate.

While the present invention has been particularly shown and described with reference to exemplary embodiments thereof, it will be understood by those of  
5 ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope of the present invention as defined by the following claims.

10

15

20